(1)

AD A120892

DATACOMPUTER AND SIP OPERATIONS

Quarterly Technical Report

January 1, 1977 to March 31, 1977

Contract No. MDA903-77-C-0183

ARPA Order No. 3346

Submitted to:

Defense Advanced Research Projects Agency
1400 Wilson Boulevard
Arlington, Virginia 22209

Attn: Dr. John Filson

Quarterly Technical Report

ARPA Order Number

3346

Name of Contractor

Computer Corporation of America

Effective Date of Contract

April 1, 1977

Contract Expiration Date

December 31, 1977

Contract Number

MDA903-77-C-0183

Principal Investigator

J. Rothnie, 617-491-3670

Short Title of Work

Datacomputer and SIP Operations

Contract Period Covered by Report

January 1, 1977 - March 31, 1977

Computer Corporation of America
575 Technology Square
Cambridge, Massachusetts 02139

Datacomputer and SIP Operations
Quarterly Technical Report
January 1, 1977 to March 31, 1977

This research is supported by the Defense Advanced Research Projects Agency of the Department of Defense under Contract No. MDA903-77-C-0183. The views and conclusior tained in this document are those of the authors and shoul be interpreted as necessarily representing the official policies, either expressed or implied, of the Defense Advanced Research Projects Agency or the U.S. Government.

Table of Contents

1. Introduction

This report describes our work during thefirst quarter of 1977 in providing very large on-line data storage and retrieval services over the Arpanet to support seismic data activity and general use. This work is being carried out under contract MDA903-77-C-0183 and represents the continuation of the operational and maintenance aspects of two previous contracts: MDA903-74-C-0225 and MDA903-74-C-0227.

These services are provided by CCA via the Datacomputer, a system designed to allow convenient and timely access by multiple remote users over a communications network to a large on-line database. The Datacomputer runs, under a modified TENEX operating system, on a DEC PDP-10 computer. This configuration has been augmented with the first public installation of the Ampex Tera-Bit Memory (TBM) System to supply the large on-line storage capacity required. Sections 2 and 3 below discuss developments in the Datacomputer and TBM.

The seismic application, the largest user of the Datacomputer, sends much of its data to CCA in real time. This real time data stream is fielded by a special dedicated processor at CCA, the Seismic Input Processor or SIP. The SIP accepts and buffers the real time data, reformats it and periodically forwards it to the Datacomputer, as further described in Section 4 below.

In order to maximize the utility of the above services to the seismic and Arpanet communities, CCA assests users in such ways as providing simple general user programs and advising users on file and retrieval design, as well as providing documentation. Activities in the area are discussed in Section 5 below.

2. The Datacomputer

This section provides an overview of the Datacomputer, a brief discussion of the two versions operational during the reporting period (Versions 3 and 3/5), and a new approach adopted in the Datacomputer, to handle a network lockup problem.

2.A. General Description

This subsection is a brief general description of the structure of the Datacomputer to provide the context for the rest of this report. Persons already familiar with the Datacomputer may safely skip it. Persons desiring a more detailed description are referred to the final Semi-Annual Technical Report for the Datacomputer Project, contract number MDA903-74-C-0225, covering July 1, 1976, through December 31, 1976.

The intended Datacomputer user is a program running on an Arpanet host remote from the Datacomputer. This program calls the Datacomputer over the network and establishes a pair of standard uni-directional 8 bit byte network connections. The user program then proceeds to send Datalanguage over one connection while the Datacomputer replies on the other.

Actually, the Datacomputer sends a "reply" to prompt the user program whenever the Datacomputer is ready to accept another line of Datalanguage. Similarly, the Datacomputer keeps the user program abreast of the progress of its requests with various synchronization, error, and success meeesages and comments. All replies have a fixed format which is designed for easy parsing by the user program. The reply begins with a unique number quickly identifying certain important messages. In the case of serious errors, to assure resynchronization with the user program, the Datacomputer enters a mode where it rejects all messages from the user, giving an appropriate reply each time, until the user program sends a special message to clear this condition.

Data as well as commands and replies can be transmitted between the user program and Datacomputer over these connections but certain characters are prohibited and the connections are fixed at an 8 bit bytesize. More general data transfers can be done by using a separate data connection.

The requests, replies, and data for a particular user program are handled initially by the half of the Datacomputer known as RH or the request handler. RH handles the parsing of the user commands and synthesis of replies. For more complex commands, RH takes the user's requests and the data descriptions stored in the Datacomputer and compiles them in several stages into code to execute the request.

Data descriptions in the Datacomputer are associated with each file of data. Data descriptions are also provided for data streams entering and leaving the Datacomputer. These streams are known as ports. Descriptions are set by the user when a file or port is created. Datacomputer data descriptions provide for hierarchical arrangements of structures of diverse data elements and repetitive lists. Many data types and data formatting alternatives are provided. This is important in ensuring that the Datacomputer can communicate with a diverse class of remote computers.

RH accomplishes many of its activities by calling on routines in the other half of the Datacomputer, known as SV or services. SV is essentially a pseudo operating system in which RH runs. It is SV that actually has custody of the Datacomputer's multi-level hierarchical directory tree and enforces the extensive protection mechanisms of the system.

A special subpart of SV, called SDAX, moves active data from slower tertiary memory to faster secondary disk storage and moves it back when secondary storage is crowded. The Datacomputer uses an Ampex TBM, as described in Section 3 below, for tertiary storage. SDAX keeps track of the location of the various copies of each file data page. It also ensures data consistency by preventing updates of a file that are not successfully completed to affect the original file.

Typically, there are multiple copies of the RH-SV pair serving multiple users over the network simultaneously. Actually, this means multiple copies of their variable areas as they are re-entrant. SDAX allows multiple readers and a single updater to be accessing the same file simultaneously. All of these users are given a consistent view of the database.

2.B. Version 3

At the start of this quarter, Version 3 became the standard operational Datacomputer. The prinicipal improvements over

Version 2 which Version 3 provides are the file groups feature and four special arithmetic functions for determining distances and directions between points on the surface of the earth.

The file groups feature provides a means of handling the large number of files involved in the seismic application. With file groups, a number of physical files with the same structure can be treated as one logical file for retrieval purposes. Furthermore, a "logical constraint" may be specified for each constituent of a group. The logical constraint indicates some restriction on the data which can exist in the file for which it is specified. For example, a constraint might require that a date field fall between 1 March 1977 and 31 March 1977. Retrieval requests against a file group will only examine those constituent files which might contain qualified data in light of the specified constraint. For example, a request for data with date fields between 1 February 1977 and 28 February 1977 would not examine any data in the file whose dates are constrained to be in March. In the seismic application, data streams are divided into a sequence of daily or monthly files. During this quarter, file groups were set up so that references against each of the streams of seismic files stored through the SIP (see section 4 below) could be made using the groups, ignoring the boundaries between the physical files.

Four sp[illegible]ial arithmetic functions were developed for the

Datacomputer in support of the ARPA Advanced Command and Control Architectural Testbed (ACCAT) project. These functions supply the great circle and rhumb line distance and bearing between two points on the earth's surface.

2.C. Version 3/5

During this quarter, some modifications were made to the Datacomputer relating to error messages and operational file flagging, and a new feature, Watch mode, was added.

As a result of the evolution of more sophisticated user programs, a number of specific error messages were assigned unique prefix codes to replace their previous default codes. The simplest user programs just send a set sequence of computed requests to the Datacomputer and give up on any failure. More sophisticated programs take different branches depending on the success or failure of various requests. More sophisticated user programs want error messages from which the particular reason for failure is easily parsible so that corrective action can be taken.

An operator command was added to the Datacomputer for flagging active files that are causing problems for the Datacomputer software or hardware due to the remnants of previous hardware problems. It is normally necessary to fix such problem files manually while the Datacomputer is not providing service to multiple users. To stop the problems from cascading, the operator can flag the problem file to prohibit all access to it until the Datacomputer can conveniently be shut down outside normal service times for repair work.

Finally, the Watch feature was added for the 3/5 Datacomputer. Using the WATCH command, a user can get reports at the beginning and end of updates by other users on a specified file. This feature was developed for the ARPA ACCAT project.

2.D. Message Lockups

Network lockups have been found to occur under unusual conditions of Datacomputer use and a solution has been devised for them. Both occur when simple user programs try to make a particular type of use of the Datacomputer.

In a typical case, a user program pays attention only to a transfer on a data connection while the transfer is generating messages to the user on the Datalanguage connection. Eventually the pipeline may fill up with messages to the user and the Datacomputer will wait until the pipeline is cleared out before providing further service to this user program. However, a simple user program will not take this action and so the job will hang up indefinitely.

An elegant and complete solution to these lockups is for user programs utilizing a separate data connection to use separate processes for the data connection and the Datalanguage connections. However, multiple processes may not be supported on the user's host computer or the user may not wish to implement them. An optional alternative has been designed which buffers messages at the Datacomputer during a transfer of data and supplies them to the user after the transfer.

3. TBM Mass Storage System

This section provides a brief overview of the Ampex TBM, the reliability problems it has caused, and a discussion of new TBM-related system software implemented by CCA during this quarter.

3.A. General Description

The CCA Datacomputer has been equipped with an Ampex Tera-Bit Memory System. This device uses video tape technology to achieve a maximum on-line capacity of 3.2 trillion bits. The current configuration at CCA supports 176 billion bits. Maximum seek time to any bit is 45 seconds. Maximum data transfer rate is a little over 5 million bits per second.

The TBM at CCA is equipped with two dual tape transport modules so at most four tapes, or about 176 billion bits (equivalent to 220 IBM type 3330 disk packs) can be available on-line. All equipment beyond the transports is non-redundant. There is one transport drive (necessary for a tape to be in motion), one data channel (necessary to encode and decode digital information to and from the broadband signal on tape),one system control processor to coordinate the TBM, and one channel interface unit that connects the TBM system to the Datacomputer's PDP-10 system. Data is transferred directly between TBM tape and core memory.

3.B. Reliability

The TBM has proved so far to be a serious obstacle to providing reliable service. The TBM has suffered failures in its control modules that exhibited a long mean time to repair (approximately four days). This significantly impaired access in January and March. Availability rates during prime operating hours (9AM to 7PM, Monday through Friday) were 78% in January, 98% in February, and 73% in March.

Problems with mechanical parts of the TBM, such as head wear on the transports, has proven to be relatively predictable and manageable through normal maintenance procedures. The difficult problems have been in the solid state control logic which might be expected to be the most reliable part of the TBM. Each of these control section problems has been solved in turn and has not re-occurred.

It seems likely that the frequency of these problems is related to CCA being the first public TBM installation and it is reasonable to hope that they will be cleared up in the near future.

3.C. TBMUTL

In the past, a collection of small and medium size TENEX programs were written for testing particular aspects of TBM operation, absolute dumping of information in readable form from TBM tape, verifying and testing correct overall TBM operation, demounting TBM tapes, reading TBM drive usage statistics, and similar utility operations. All these programs were written in machine language to meet particular needs.

This quarter a unified TBM utility program, called TBMUTL, was developed to replace all of the earlier programs. For ease in development and future maintenance and expansion, TBMUTL was written in BCPL, a high level system implementation language.

3.D. CCA TENEX

The TBM handling routines in CCA's TENEX system were modified to provide more sophisticated handling of multiple errors and more capabilities in using the TBM.

For TBM error recovery, it is necessary to decide how many times to retry an operation in the face of errors which may be transient. Based on our error experience, the number of attempts for retryable errors was reduced to avoid excessively degrading system throughput in the rare cases where irrecoverable errors are being encountered. The criterion for declaring a TBM drive down pending operator action was made considerably more complex. Errors were divided into two categories, fatal and suspicious. Fatal errors cause the drive to be declared down immediately. Suspicious errors, which do not include any errors recovered by retries, are counted and the drive declared down after sixteen. Warning messages for a TBM drive are also counted and a drive declared down after a large number of them. When a tape is mounted, the error and warning counts are cleared.

TENEX was also augmented with a system call to issue a read recover command to the TBM. Read recover, which Ampex is in the process of implementing in the TBM internal software, will automatically perform all the normal manual steps taken to recover a block of data for which difficulties in reading are being encountered.

4. The SIP

The SIP, or Seismic Input Processor, matches the real time continuous seismic data stream arriving at CCA to the Datacomputer. It is a small dedicated system implemented on a DEC PDP-40 computer with two RP04 disk drives and an Arpanet interface.

4.A. Reliability

The SIP was operational for all of this quarter receiving data from the CCP, buffering it on its disks, and reformatting it and forwarding it to the Datacomputer. Two problems were encountered. In one, due to problems at the PLURIBUS IMP at SDAC, very rapid bursts of type 6 and 7 Arpanet messages overflowed a queue in the SIP. In the other, a peculiar set of circumstances that had not been encountered before resulted in a buffer lock-up. Minor changes were made in the SIP to prevent a re-occurrence of these problems.

To assist in SIP operations, a new complete operator's manual was distributed to replace the former preliminary manual. This new manual includes instructions for changing disk packs. This operation has been necessary occasionally when TBM hardware problems have made the Datacomputer unavailable for more than two days.

4.B. Arpanet Considerations

Difficulties have been encountered in using the Arpanet for the continuous high bandwidth transmission of seismic data. It has been determined that the current protocol used for the real time data does not make the most efficient use of the network and that some of the problems are due to lack of sufficient reassembly buffer space in the CCA IMP and competition between the real time seismic and other traffic.

Two steps are being taken to alleviate these problems. First, a PLURIBUS IMP is scheduled to be installed at CCA in July to provide ample network reassembly buffers. Second, a new protocol has been designed for real time seismic data transmission that tries to minimize the number of separate messages by sending maximum length messages. The new protocol should approach the maximum possible efficiency of network utilization.

5. User Coordination

CCA's user coordination work during the past quarter fell into three categories, work on general purpose user programs, facilities and changes to operationally improve the Datacomputer, and assistance to and investigations on behalf of users.

5.A. User Programs

Subroutine interfaces to the Datacomputer were made available for COBOL through a package called DCPKG and for BCPL through an interface to the previously existing TENEX machine language DCSUBR package.

DFTP, a program for simple file storage on the Datacomputer, was adapted to changes in the Version 3 Datacomputer. Data stored with DFTP, which is used by more than 170 users on the Arpanet, increased from about 1300 to 2228 megabits.

5.B. Operational Coordination

Four steps were taken this quarter to improve operational aspects of Datacomputer use.

First, all preventive maintenance activity on the TBM and TENEX hardware underlying the Datacomputer was scheduled to occur before 9 A.M. 9 A.M. to 7 P.M. weekdays was established as our prime time. All reasonable efforts are made to keep the

the Datacomputer up and providing service on the network during this period.

Second, a rotating position of Programmer of the Week was established within CCA. The primary duty of the POW is to see that user enquiries and complaints are responded to.

Third, an answering service (telephone number 617-482-6226) and a telephone company beeper were obtained so that urgent operational messages can get through to a CCA employee even when the Datacomputer is unattended. The message is given to the answering service who will beep the CCA person who can then call in and get the message.

Fourth, the automatic Datacomputer status facility provided at CCA was augmented. This facility can be called over the network and will tell whether or not the Datacomputer exists on the CCA TENEX system, if it is listening for new users, and the state of its network connections.

5.C. Investigations

Through the Programmer of the Week and otherwise, CCA continued to provide assistance to the Datacomputer user community.

Two particular in depth investigations were made this quarter for the seismic community. The first of these was to determine the optimum way to extract non-array long period data between two arbitrary seconds. The second was to find the optimum way to extract detection start and end times from large non-array short period data files.